知识管理论坛 ISSN 2095-5472 CN11-6036/C
Knowledge Management Forum

E-mail: kmf@mail.las.ac.cn
http://www.kmf.ac.cn

【学术探索】

语义联合目录：基于 BIBFRAME 2.0 的联合目录建构

◎ 林泽斐

福建师范大学社会发展学院 福州 350013

摘要：[**目的 / 意义**] 随着以 BIBFRAME 为代表的语义书目模型不断完善，书目数据面临着从 MARC 向语义化书目的转型。本文探索了基于语义化书目数据的联合目录构建模式。[**方法 / 过程**] 首先分析传统联合目录的建设模式，认为传统联合目录建构方法在语义网环境下有一定借鉴意义；其次在分析 BIBFRAME 2.0 书目信息表示形式的基础上，提出 3 种语义联合目录构建模式：集中式语义联合目录、分布式语义联合目录和集中与分布相结合的语义联合目录；最后利用 Jena Fuseki+Apache HTTP 服务器搭建实验平台，对语义联合目录的互操作模式进行测试。[**结果 / 结论**] 实验结果显示，本研究提出的 3 种语义联合目录建设方法在实践上具有可行性，可为图书情报部门的语义化书目数据互操作和跨库检索系统的实现提供有益参考。

关键词：联合目录 BIBFRAME 语义网 关联数据

分类号：G257.22

引用格式：林泽斐 . 语义联合目录：基于 BIBFRAME 2.0 的联合目录建构 [J/OL]. 知识管理论坛 , 2016, 1(6): 440-448[引用日期]. http://www.kmf.ac.cn/p/1/77/.

自 1966 年美国国会图书馆制定机读目录规范以来，MARC 已成为书目数据的主流交换格式，同时也是当前联合目录构建的基石。尽管 MARC 在图书情报领域得到了广泛应用，但学术界长期存在着对 MARC 的批评，如 L. Andresen 认为 MARC 缺乏灵活性，导致对资源的附加描述成为本地系统和联机系统交流的障碍 [1]；G. Alemu 等认为 MARC 在设计之初是面向文献的，这使得 MARC 难以描述非文献实体之间的关系 [2]；R. Tennant 在著名的 MARC Must Die 一文中则指出 MARC 的使用囿于图书馆领域，且 MARC 数据过于平面化，难以表达实体间的等级关系 [3]。

在近 20 年时间里，美国国会图书馆、OCLC 等组织对 MARC 进行了诸多改进，发展出 MARC21、MARCXML 等格式；在 RDA 发布后，图书馆界也迅速对 MARC21 中的相关字段进行了增补和修改，以便与 RDA 保持一致。然而从整体上看，MARC 以文档为中心的设计思想已成为自身发展的禁锢，简单地将 MARC 字段映射为 XML 也并无法充分揭示文献所涉及的实体与关系。针对 MARC 存在的缺陷，图书

作者简介：林泽斐（ORCID：0000-0001-8637-7359），讲师，硕士，E-mail: linzf@fjnu.edu.cn。

收稿日期：2016-08-04 发表日期：2016-12-23 本文责任编辑：易飞

知识管理论坛，2016(6):440–448
DOI：10.13266/j.issn.2095-5472.2016.051

馆界开始探索利用关联数据进行书目元数据的描述，试图建立更加开放、有足够伸缩性，且能够溶入语义网环境的文献数据交换标准。以BIBFRAME（Bibliographic Framework）为代表的语义书目模型，正是图书馆界在替代MARC道路上迈出的重要一步。虽然BIBFRAME仍在完善过程中，但鉴于国会图书馆在书目数据领域的影响力，BIBFRAME将不可避免地对未来的联合目录构建形式产生重要影响。

1 BIBFRAME及其研究概况

BIBFRAME即书目框架，是美国国会图书馆联合以语义网技术见长的数据服务公司Zepheira共同研发的书目数据模型，旨在使用关联数据对书目数据进行描述，从而实现对MARC的替代。国会图书馆于2012年底发布了BIBFRAME草案，并于2014年公布了BIBFRAME正式版本。BIBFRAME正式发布后，国会图书馆根据BIBFRAME应用过程中的反馈，于2016年4月发布了经修改后的BIBFRAME 2.0模型和词汇表。

在BIBFRAME 1.0模型中，包括子类和子属性在内，共包含53个类及289个属性[4]，其中Work、Instance、Authority和Annotation四个类为BIBFRAME的核心类：作品（Work）用于反映作品的概念实质；实例(Instance)反映了作品的物理载体表现形式，一个作品可以有多个对应的实例，但一个实例只有一个对应的作品；规范（Authority）实现了与作品和实例相关的人物、组织、事件、主题的规范控制；注释（Annotation）则提供有关作品和实例相关的描述信息（如封面、描述、评论、馆藏等）。

BIBFRAME 2.0对BIBFRAME 1.0的体系结构进行了优化。包括子类和子属性在内，BIBFRAME 2.0共设置了140个类及165个属性[5]。BIBFRAME 2.0中属性减少而类增加的主要原因在于：BIBFRAME 1.0中，对于一种资源的多种类型，使用多个属性来表示，如表示不同类型Identifiers（标识符）的属性包括bf:isbn、bf:issn、bf:lccn等；而在BIBFRAME 2.0中，相关标识符均使用bf:identifiedBy属性表示，标识符的类型则反映为bf:Isbn、bf:Issn、bf:Lccn等多个类。

BIBFRAME 2.0的另一重要变化是将4个核心类简化为3个：Work、Instance和Item。其中Work、Instance类的含义与1.0版相同；Item（馆藏项）类代表一个实例的馆藏复本（可以为实体馆藏或电子馆藏），并反映了诸如馆藏位置、条码号、索书号等与馆藏复本相关的信息，其内涵与BIBFRAME 1.0中Annotation类的HeldItem子类相似。图1是采用RDF/Turtle格式进行形式化表述的BIBFRAME 2.0馆藏书目数据片段。该数据包含一个作品（bf:Work）、一个作品的实例（bf:Instance）和该实例的馆藏项（bf:Item），并分别定义了题名（bf:title）、作者（bf:contributor）、ISBN号（bf:Isbn）、所在图书馆（bf:heldBy）和条码号（bf:Barcode）等数据项。

BIBFRAME出现后，部分国外机构和学者开展了相关试验与研究。截至2016年4月，共有7个机构（美国国会图书馆、不列颠图书馆、德国国家图书馆、普林斯顿大学图书馆、乔治·华盛顿大学图书馆、美国国家医学图书馆、OCLC）在BIBFRAME项目平台中发布了书目测试数据集[6]。欧洲数字图书馆（Europeana）项目的研究人员探索了欧洲数据模型（Europeana Data Model，EDM）与BIBFRAME的映射关系[7]。康奈尔大学图书馆、哈佛大学图书馆和斯坦福大学图书馆的合作项目Linked Data for Libraries（LD4L）则试图通过整合包括BIBFRAME、VIVO、VIAF和ISNI在内的关联数据集，创建一个面向学术资源的语义信息存储模型[8]。

我国图书馆学界对BIBFRAME的研究始于2014年。刘炜、夏翠娟率先撰文对BIBFRAME的体系结构、特性和应用领域进行了探讨[9]。笔者通过对中文数据库进行检索发

现，截至 2016 年 4 月，我国针对 BIBFRAME 的学术论文共有 11 篇，研究主要着眼于 3 个方面：一是 BIBFRAME 的体系架构和应用领域；二是 MARC（CNMARC、MARC21）向 BIBFRAME 转换路径；三是 BIBFRAME 与 FRBR、RDA、Shema.org 等相关模型和规则的比较研究。目前国内有关 BIBFRAME 的应用研究还较为缺乏，夏翠娟等的《基于书目框架（BIBFRAME）的家谱本体设计》一文是唯一对 BIBFRAME在具体应用领域进行研究的论述[10]，有关基于 BIBFRAME 模型的联合目录构建模式未见有研究涉足。

```
@prefix bf: <http://id.loc.gov/ontologies/bibframe/>.
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>.
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>.
<http://library.org/resources/work/9406476> a bf:Text, bf:Work;
bf:title [ a bf:WorkTitle;
rdfs:label “Across the river and into the trees” ];
bf:contributor<http://viaf.org/viaf/97006051>.
<http://library.org/resources/instance/5866658> a bf:Instance;
bf:instanceOf<http://library.org/resources/work/94064763>;
bf:identifiedBy [ a bf:Isbn; rdf:value “9780743565172” ].
<http://library.org/resources/item/72616221> a bf:Item;
bf:itemOf<http://library.org/resources/instance/5866658>;
bf:heldBy<http://library.org/organization/sample-lib>;
bf:identifiedBy [ a bf:Barcode;rdf:value “001937362” ].
```

图 1　BIBFRAME 2.0 馆藏书目数据实例

2 传统联合目录建设模式分析

联合目录（union catalogs）在联合检索、联机编目、馆际互借等馆际资源共建共享业务中具有重要作用。联合目录最早始于 1898 年普鲁士 10 所大学图书馆编制的《柏林印刷本目录》，随后在美国国会图书馆等机构的推动下，逐步发展出卡片式联合目录、缩微胶片式联合目录、光盘版联合目录和联机目录等多种形式[11]，其中联机目录是当前联合目录建设的主流形式。从体系结构上看，传统联机目录总体可分为两种类型：集中式联机联合目录和分散式联机联合目录。

2.1 集中式联机联合目录

集中式联机联合目录的典型特征为存在一个中央书目数据库，该数据库通过一定机制与成员馆书目数据实现同步更新。集中式联合目录实现的关键在于中央数据库与成员馆的数据同步机制。当前联机目录的建设实践主要基于人工批量同步和元数据收割两种形式。如我国台湾地区的“全国图书书目资讯网”（NBINet）即是前者的代表。根据 NBINet 的《书目网络合作办法》，NBINet 成员馆须定期或不定期向台湾“国家图书馆”寄送 CMARK、USMARK 或 MARK21 格式的档案，“国家图书馆”则负责将成员馆的 MARC 数据导入 NBINet 数据库[12]。CALIS 联合目录数据库也采用了集中式数据库结构，在管理中心和地区中心建立 8 个联合目录数据库，通过“增量复制”技术（互为“镜像”）使 8 个联合目录数据库中的书目记录保持准同步[11]。基于元数据收割的联机目录多采用 OAI-PMH 协议。

OAI-PMH 定义了两类角色：数据提供者（data prodiver）和服务提供者（service provider）。数据提供者将一个或多个书目数据仓储以 OAI 协议进行发布，服务提供者则使用

知识管理论坛，2016(6):440–448
DOI：10.13266/j.issn.2095–5472.2016.051

支持 OAI 协议的客户端对 OAI 仓储进行定期收割。当前在 openarchives.org 中公布的数据提供者已达 3 014 个 [13]。国内的 CALIS 高校学位论文数据库和中国科学院联合机构知识库也采用了成员馆分散建库、中心馆利用 OAI-PMH 集中收割的建设模式 [14]。

2.2 分布式联机联合目录

分布式联机联合目录由不同文献机构维护各自的书目数据库，书目检索方通过联邦检索的形式向多个文献机构广播检索请求，并将返回的检索结果进行汇总。当前主流的分布式联机目录大多基于 Z39.50 协议，即检索客户端将检索请求转换为 Z39.50 协议的应用协议数据单元（APDU）后广播至多个支持 Z39.50 协议的服务器。服务器对 APDU 解码后，转换成自身系统的检索请求进行查询。查询完成后，使用相同的方法将检索结果返回给检索客户端。如香港科技大学的《香港大学图书馆联合目录》就利用 Z39.50 实现了包括 8 所香港大学和 16 个国外机构在内的文献资源分布式检索 [15]。

3 联合目录的语义化构建

语义网是 T. Berners-Lee 于 1998 年所提出的概念，其目标是为万维网中的文档添加可理解的语义；2006 年，T. Berners-Lee 在语义网的基础上提出了关联数据（linked data）一词，试图利用 Web 实现对 RDF 数据的相互关联，最终形成数据之网（Web of Data）。语义网和关联数据的概念提出后，图书馆界也在不断探索书目数据语义化的策略。目前图书馆界的数据语义化尝试主要集中在两个方面：一是词表和分类表数据的语义化，如美国国会图书馆已将《美国国会图书馆标题词表》（Library of Congress Subject Headings）进行语义化描述（基于 SKOS）后以关联数据形式发布在 Web 上，台湾“国家图书馆”也已完成了主题词表的 SKOS 化和中文图书分类法的语义化转换，并将其应用于其编目系统 [16]；二是书目数据集的语义化，如大英图书馆将英国国家书目（BNB）从 MARC21 格式转换为关联数据的 RDF/XML 格式，并通过 Web 提供数据集的下载；德国柏林自由大学和汉诺威大学的研究者则利用 D2R 服务器将著名的计算机科技文献书目数据库 DBLP 发布为关联数据 [17]。

近年来，随着关联数据在图书情报领域应用的深化，部分机构已开始尝试实现联合目录的语义化构建。最早将联合目录以语义化形式发布的机构为瑞典国家图书馆，该馆于 2008 年尝试将瑞典联合目录（Swedish Union Catalogue）发布为关联数据 [18]。作为全球最大联机目录，OCLC 的 WorldCat 数据库自 2012 年开始将一批被广泛收藏的作品通过 Schema.org 词汇表加以描述，并以关联数据的形式进行发布，目前已发布书目数据 1.97 亿条 [19]。

从当前语义化联合目录的建设实践上看，主流构建方式为将机构原先拥有的 MARC 数据内容进行语义网转换，通过机构自身所编制的书目本体对书目进行语义化描述，最终以 Turtle、RDF/XML、RDFa 等 RDF 形式化表示进行发布。这种语义化联合目录构建方式的不足之处是无法实现原生的语义联合目录构建；联合目录与成员馆之间缺乏语义化的数据互操作，导致语义书目数据同步的滞后性；同时也缺乏基于以 BIBFRAME 为代表的标准化语义书目数据格式的联合目录构建实践。因此，如何构建 BIBFRAME 作为标准的书目数据载体并能够实现原生语义化互操作的联合目录，是书目数据语义化过程中亟待解决的问题。

4 基于 BIBFRAME 2.0 的语义联合目录体系结构

如前所述，传统联机目录可采用集中式和分散式两种建构策略。笔者认为，传统联机目录的建构方法在语义网环境下有一定借鉴意义，但不可简单套用。构建基于 BIBFRAME 2.0 的语义联合目录，可利用 BIBFRAME 的开放性及跨域关联特性，并借鉴传统联机目录中 OAI 收割机制及基于 Z39.50 的广播式查询机制，采



知识管理论坛
2016 年 第 6 期(总第 6 期)

用集中式、分布式和集中与分布相结合 3 种构建模式。

4.1 集中式语义联机目录

集中式语义联合目录的特征为存在一个中心语义数据仓储。如图 2 所示，中心语义数据仓储中存放有多个数据集，分别用于存储使用 BIBFRAME 2.0 表示的作品、实例和馆藏项三元组以及各成员馆、书库、作者的 URI 命名等规范数据。所有数据集使用带有访问权限控制的 SPARQL 端点（Endpoint）将数据开放给成员馆。各成员馆亦分别维护一份本地 BIBFRAME 数据仓储，同样使用 SPARQL 端点将数据开放给中心仓储。

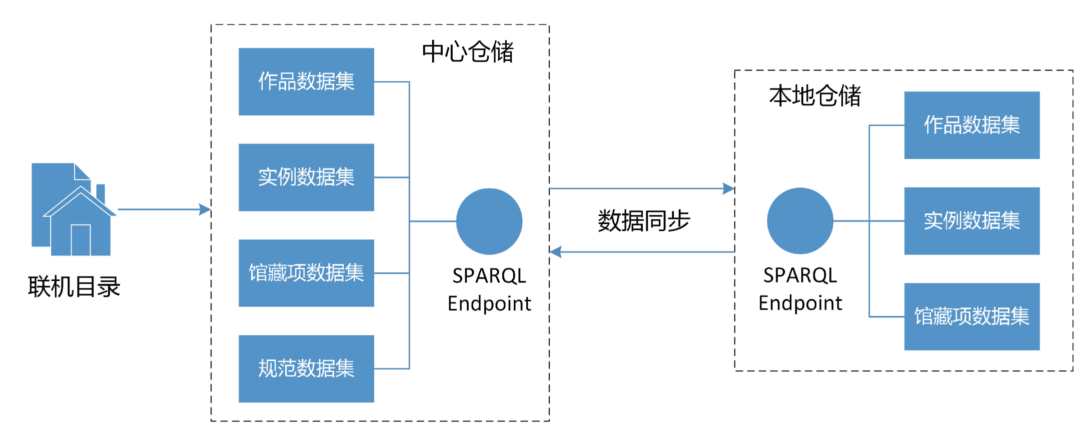


图 2 集中式语义联机目录结构

集中式语义联合目录实现的关键在于中心仓储数据集与成员馆本地数据集的同步机制。数据集同步可采用两种形式：

（1）批量同步。即中心系统轮询成员馆的 SPARQL 端点，定期将本地书目仓储的数据收割至中心仓储。批量同步方式是 OAI-PMH 收割在语义网环境下的替代。

（2）增量同步。增量同步将语义联合目录构建与联机编目相结合，其运作流程为：成员馆获得新馆藏资源时，首先查询本地仓储是否有相应作品、实例和馆藏信息，若无则通过 HTTPPOST 方法将 SPARQL 请求发送至中心仓储的 SPARQL 端点进行检索。若中心仓储已存在该资源的相应数据，则成员馆将其套录至本地仓储中；否则，成员馆完成相关数据的原始编目并同步存储至本地和中心仓储。为防止数据更新后导致的不一致，可在中心系统和本地系统间建立 Pingback 机制，若中心仓储中的数据发生了修改，则各成员馆本地系统将收到相应书目的更新通知。增量同步的优势在于可有效防止元数据的重复建设，因此特别适合作品（Work）和实例（Instance）等共性较强的数据的同步。

4.2 分布式语义联合目录

分布式语义联合目录不设置中心语义数据仓储，而是将 BIBFRAME 数据分散存储于不同机构中。各机构分别承担自身收藏的所有书目的作品、实例及馆藏项数据的建设，并将 BIBFRAME 数据以 SPARQL 端点的形式向外部开放。

当联合目录收到用户检索请求时，可分别向各成员馆发送 SPARQL 查询请求，再将响应的结果集进行合并，也可利用 SPARQL 的联邦查询（federated query）机制，一次性对指定机构的书目进行联邦查询，并将检索结果进行去重后提供给用户。图 3 为一段 SPARQL 查询实例，该实例通过联邦检索方式查询了指定 ISBN 号的图书实例在两馆中各自的馆藏复本量。

分布式语义联合目录无需成员馆对系统架构作较大调整，也无需部署专用的同步终端软

知识管理论坛，2016(6):440–448
DOI：10.13266/j.issn.2095–5472.2016.051

件，因此较适合松散型图书馆联盟成员间的资源整合。其缺点是不同成员间缺乏协调机制，可能存在对同一作品和实例数据的重复建设，并由此产生数据去重过程中的取舍问题。

```
SELECT ?lib (COUNT(*) as ?itemNum) WHERE {
{
SERVICE <http://library1.org/sparql> {
?instance bf:identifiedBy ?isbn.
?isbnrdf:value "9780743565172".
?item bf:itemOf ?instance.
?item bf:heldBy ?lib. }
}UNION {
SERVICE <http://library2.org/sparql> {
    ?instance bf:identifiedBy ?isbn.
    ?isbnrdf:value "9780743565172".
    ?item bf:itemOf ?instance.
?item bf:heldBy ?lib. }
}
} GROUP BY ?lib
```

图 3 基于联邦查询的分布式书目检索

4.3 集中与分布相结合的语义联合目录

集中与分布相结合的语义联合目录综合了前两者的特性。在这一模式下，作品、实例和命名规范等共性较强的数据采用集中式管理；馆藏项等个性较强的数据采用分布式管理（见图 4）。其优点在于通过作品和实例数据的集中管理可以减少资源的重复建设，同时通过分布式的采集馆藏项数据可以尽可能扩大检索的成员馆范围。

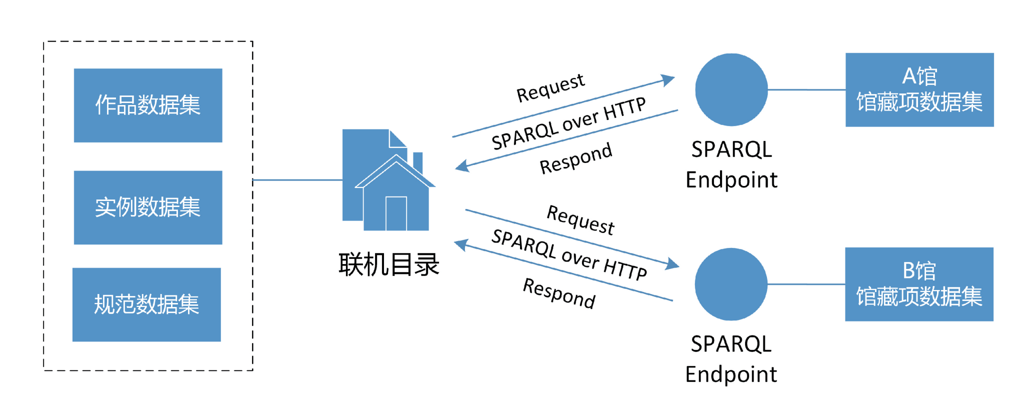


图 4 集中与分布相结合的语义联合目录结构

5 实验测评

为验证上述语义联合目录构建方法的有效性，笔者使用 Jena Fuseki+Apache HTTP Server 搭建实验平台，进行了相关测试。Fuseki 是 Apache 基金会开发的开源语义网框架 Jena 中的一个 SPARQL 服务器，其内置了 TDB 三元组存储器，同时提供支持 HTTP REST 架构的 SPARQL 端点（Endpoint）服务。

Fuseki 服务器自带有基于 Apache Shiro 框架

的权限控制机制[20]，但经笔者测试，其权限控制主要针对三元组数据集管理系统，并未覆盖 SPARQL 端点服务，这使得 SPARQL 端点暴露在外，存在一定的数据安全隐患。为解决这一问题，笔者利用 Apache HTTP Server 的反向代理机制构建了反向代理服务器指向 Fuseki 服务器的 URL，并针对该反向代理设置了 AuthType Basic 权限限制，从而将 SPARQL 端点覆盖在鉴权范围内。

针对集中式语义联合目录的馆际互操作问题，笔者按上述权限控制方案在局域网中搭建了两台部署有 Fuseki+Apache HTTP Server 的服务器，用于模拟中心仓储和本地仓储，并在中心仓储服务器中建立了名为 center 的数据集（Dataset），在本地仓储服务器中建立了名为 local 的数据集，两数据集中均建立了名为 work、instance、item 的 3 个具名图（Named Graph），用于分类保存用于测试的 BIBFRAME 2.0 RDF 数据。

Jena ARQ 是 Jena 所包含的 SPARQL 查询引擎，提供了包括基本 SPARQL 查询、联邦查询及 SPARQL 更新在内的一系列 API。笔者利用 Jena ARQ API 实现了对远端语义书目仓储的查询和修改，进而可完成数据批量下载、套录等一系列操作。图 5 为书目数据读写操作的关键代码（其中 admin 和 pw 为鉴权信息，http://192.168.1.2/center/sparql 和 http://192.168.1.2/center/update 分别为经反向代理包装后的中心仓储的查询和修改 SPARQL 端点）：

```
/* 数据查询 */
String queryString = “select * where { graph <http://192.168.1.2/center/work>
                              { <http://lib.org/work/98767> ?p ?o. } }”; // 查询作品
Query query = QueryFactory.create(queryString);
HttpAuthenticatorauth = new SimpleAuthenticator(“admin”,”pw”.toCharArray());
QueryExecutionqexec = QueryExecutionFactory.sparqlService(
                              “http://192.168.1.2/center/sparql”, query, auth);
ResultSet results = qexec.execSelect();
while (results.hasNext()) { ... }  // 迭代读取结果集
/* 数据修改 */
SimpleAuthenticator(“admin”,”pw”.toCharArray());
String updateString = “prefix bf: <http://id.loc.gov/ontologies/bibframe/>”+
“insert data { graph <http://192.168.1.2/center/work>”+
                      “{ <http://lib.org/work/98767> a bf:Work. } }”; // 添加作品数据
UpdateRequest update = UpdateFactory.create(updateString);
UpdateProcessRemoteupdateProcess = (UpdateProcessRemote) UpdateExecutionFactory
                      .createRemote(update, “http://127.0.0.1:3030/libraryB/
update”);
updateProcess.setAuthenticator(auth);
updateProcess.execute();
```

图 5 基于 Jena ARQ API 的书目数据读写

为验证分布式语义联合目录互操作的可靠性，笔者使用 Java 程序自动生成了 1 000 条 BIBFRAME 作品测试数据，每部作品数据各生成 10 条相关实例，每部实例各生成 10 个相关馆藏项，共计 10 万条馆藏数据，并将上述数据分散存放于广域网中的 3 台计算机中（其中 2 台位于 Chinanet，1 台位于 Cernet）。此后，笔者以 Chinanet 中的另一主机充当联合目录服务

知识管理论坛，2016(6):440–448
DOI：10.13266/j.issn.2095–5472.2016.051

器对上述 3 台计算机进行 SPARQL 查询，以联邦检索方式查询特定作品的所有馆藏信息。实验结果显示，10 次针对不同作品的查询均得到正确的查询结果，取得返回结果所用的平均耗时为 1.2 秒，但实验也发现基于 SPARQL 联邦查询获取数据存在一定的"木桶效应"，即一旦某一节点连接时延过长或断开连接，将导致联合目录查询超时。因此，笔者认为基于联邦检索方式实现的分布式语义联合目录主要适用于网络时延较短、成员馆节点数量较少的应用场景。在网络时延不确定或需要大规模分布式查询的应用场景中，通过前述 Jena ARQ API 分别查询成员馆 SPARQL 端点，再进行数据合并的方式应更为理想。

6 总结与展望

基于 BIBFRAME 2.0 的语义联合目录相比传统联合目录相比，主要优势在于：

（1）实现了书目信息内容揭示和馆藏数据的分离。在 BIBFRAME 模型中，作品数据属于抽象内容，实例和馆藏数据属于载体表现，在描述时分开描述。在语义联机目录中，作品、作品实例和馆藏项三元组可以分别存储于不同数据集中，由不同机构或部门负责维护，用户在联机检索时再整合多个数据集的内容进行查询。

（2）提高了书目数据的开放性和关联程度。在语义网中，使用 HTTP URI 作为实体的标识符。在 BIBFRAME 书目数据中，任何与图书有关的实体（如作者、出版商、图书馆、地区）均可使用 URI 进行标识。这有助于打破传统图书馆的封闭系统与万维网之间的藩篱，便于实现书目数据和其他开放数据集之间的混搭。

针对 BIBFRAME 2.0 出现后书目数据语义化的发展，本研究在分析传统联合目录构建机制的基础上提出了 3 种语义联合目录的构建模式，即集中式语义联合目录、分布式语义联合目录以及集中与分布相结合的建设模式，通过语义联合目录的构建，可实现馆际间语义书目数据的联合查询，进而可为基于 BIBFRAME 的语义联机编目和馆际互借等馆际协作提供基础。

为验证语义联合目录构建方法的有效性，本研究开展了基于 BIBFRAME 和语义仓储的馆际互操作实验。由实验结果可知，本研究提出的语义联合目录建设方法在实践中具有可行性。在本研究中，主要以自动生成的书目数据作为测试资源，未将大批量 MARC 数据转换后进行测试，同时也缺乏对大量成员馆节点在跨区域网络环境下构建联合目录的测试。在后续研究中，笔者将对 CNMARC 向 BIBFRAME 2.0 的转换机制进行探索，同时对跨区域多节点环境下的语义联合目录应用予以关注和测试。

参考文献：

[1] ANDRESEN L. After MARC-what then[J]. Library hi tech, 2004, 22(1): 40-51.

[2] ALEMU G, Stevens B, Ross P, et al. Linked data for libraries: benefits of a conceptual shift from library specific record structures to RDF-based data models[J]. New library world, 2012, 113(11/12): 549-570.

[3] TENNANT R. MARC must die[J]. Library journal, 2002, 127(17): 26-28.

[4] Library of Congress. BIBFRAME vocabulary[EB/OL]. [2016-04-10]. http: //bibframe.org/vocab-list.

[5] Library of Congress. BIBFRAME 2. 0 vocabulary[EB/OL]. [2016-04-15]. http: //id.loc.gov/ontologies/bibframe.html.

[6] Library of Congress. BIBFRAME Implementation & testing[EB/OL]. [2016-04-10]. http://www.loc.gov/bibframe/implementation.

[7] Zapounidou S, Sfakakis M, Papatheodorou C. Library data integration: towards BIBFRAME mapping to EDM[C]// Closs S, Studer R, Garoufallou E, et al. Metadata and Semantics Research. Berlin: Springer International Publishing, 2014: 262-273.

[8] LD4L Project team. Project description [EB/OL]. [2016-04-30]. https://www.ld4l.org/description.

[9] 刘炜 , 夏翠娟 . 书目数据新格式 BIBFRAME 及其应用 [J]. 大学图书馆学报 , 2014(1): 5-13.

[10] 夏翠娟 , 刘炜 , 张磊， 等 . 基于书目框架 (BIBFRAME) 的家谱本体设计 [J]. 图书馆论坛 , 2014(11): 5-19.

[11] 卢共平 , 汪善建 . 欧美国家联合目录的进展与我国虚拟联合目录的发展思路 [J]. 图书情报工作 , 2002, 46(10): 84-87.

[12] NBINet. 全国图书书目资讯网合作编目要点 [EB/OL]. [2016-04-30]. http: //nbinet.ncl.edu.tw/content.



aspx?t=m&id=95.
[13] Open Archives Initiative. Registered data providers[EB/OL]. [2016-04-30]. http://www.openarchives.org/Register/BrowseSites.
[14] 姚晓娜 , 祝忠明 , 卢利农，等 . 机构知识库 OAI 互操作数据同步策略研究 [J]. 现代图书情报技术 , 2014(3): 14-18.
[15] Hong Kong University of Science and Technology Library. Z39. 50 union catalogs [EB/OL]. [2016-05-03]. http://ustlib.ust.hk:211/z39m.
[16] 台湾“国家图书馆”．“国家图书馆”链接资源系统开放使用 [EB/OL]. [2016-10-23]. http://catweb.ncl.edu.tw/portal_d2_page.php?button_num=d2&cnt_id=301.
[17] 欧石燕 . 语义网的主要功能及其在数字图书馆中的应用 [J]. 数字图书馆论坛 , 2014(3): 2-10.
[18] Martin M. Making a library catalogue part of the semantic Web[EB/OL]. [2016-05-15]. http://dcpapers.dublincore.org/pubs/article/view/927/923.
[19] OCLC. Data strategy and linked data[EB/OL]. [2016-05-03]. http://www.oclc.org/data.en.html.
[20] Apache Software Foundation. Security in Fuseki2[EB/OL]. [2016-05-15]. https://jena.apache.org/documentation/fuseki2/fuseki-security.html.

Semantic Union Catalogs: The Construction of Union Catalogs Based on BIBFRAME 2.0

Lin Zefei
College of Social Development, Fujian Normal University, Fuzhou 350013

Abstract: [Purpose/significance] With the continuous improvement of bibliographic description models such as BIBFRAME, bibliographic data are facing the transition from MARC to the semantic bibliography. This paper explores the construction model of semantic union catalogs. **[Method/process]** The author analyzed the construction model of tranditional union catalogs firstly, and considered that its pattern had reference significance for semantic union catalogs but should not be simply copied. Then, the authors analyzed the bibliographic description method of BIBFRAME 2.0, and proposed three construction models of semantic union catalogs: the centralized semantic union catalogs, the distributed semantic union catalogs and the combination of these two methods. Finally, the authors built an experimental platform based on Jena Fuseki and Apache HTTP server, and tested the interoperated method of semantic union catalogs. **[Result/conclusion]** The experimental results show that the three methods proposed in this paper are feasible in practice, and can provide useful references for the semantic bibliographic data interoperability and the cross-database retrieval system of the library and information department.

Keywords: union catalogs　BIBFRAME　Semantic Web　linked data